

Reprint from

Current Topics in Microbiology and Immunology, Vol. 240
Plant Biotechnology: New Products and Applications

Editors: J. Hammond, P. McGarvey, and V. Yusibov

© Springer-Verlag Berlin Heidelberg 1999

Printed in Germany. Not for Sale.

Springer

Transgenic Plants for Therapeutic Proteins: Linking Upstream and Downstream Strategies

C.L. COOPER¹, J.G. BOONER², and K.K. OSMI³

1	Introduction	93
2	Plant-Based Biopharmaceutical Production: Issues and Answers	94
2.1	Selection of Crop Species	100
2.2	Order of Trans	100
2.3	Expression Strategies	101
2.4	Posttranslational Processing	103
2.5	Antibody Production	105
2.6	Antibody Purification	106
2.7	Antibody Formulation	107
3	Examples of Plant-Synthesized Protein Therapeutics Linking Upstream and Downstream Strategies	108
3.1	Production of Human Cystic Fibrosis Transmembrane Conductance Regulator	109
3.2	Production of Human Interferon α in <i>Arabidopsis</i>	112
3.3	Production of Human Interferon γ in <i>Arabidopsis</i>	113
4	Summary	114
	References	115

1 Introduction

With the new knowledge generated through the Human Genome Project and related biomedical research comes a potential revolution in drug development strategies. One of the most direct applications of this knowledge will be highly specialized recombinant protein-based therapeutics. Recombinant drugs such as human erythropoietin (EPO), tissue plasminogen activator (tPA), and Charysin[®] (glucosylated) are currently on the market and many other recombinant proteins are in various stages of human clinical trials. Commercial production of

¹ CropTech Corp., Virginia Tech Corporate Research Center, Blacksburg, VA, 24060 and Department of Plant, Soil, and Environmental Science, Virginia Polytechnic Institute and State University, Blacksburg, VA, 24061-0346, USA.
² Sanofi-Schering-Plough, 2000 North 1st St., P.O. Box 1518, Kenilworth, NJ, 07033, USA.
³ CropTech Corp., Virginia Tech Corporate Research Center, Blacksburg, VA, 24060, USA

these proteins utilizes fermentation (primarily *E. coli* and yeast) and mammalian cell systems (e.g., Chinese hamster ovary cells), the major expression systems adopted by the well established biotechnology companies. However, these expression systems have significant limitations. Bacteria cannot perform the complete posttranslational modifications required for bioactivity of many human proteins and high-level expression often leads to accumulation of insoluble protein aggregates. While mammalian cell cultures perform the required protein modifications, low transgene expression levels, instability of selected cell lines, and the difficulties and high expense of scale-up are often limiting or severely impact cost. Thus, there remains significant opportunity for alternative expression systems that address these limitations and cost issues to compete in the protein therapeutics market. In fact, development of more cost-effective protein bioproduction systems may be critical in realizing the discoveries of genomics and medical research into widely available and affordable treatments and cures. Recent advances in the area of transgenics—the use of genetically engineered plants and animals for bioproduction—indicate great promise as effective protein factories. The fact that recombinant proteins from both transgenic animals and transgenic plants are now in clinical trials demonstrates significant progress toward commercialization of these technologies.

For any particular target protein, selection of a recombinant system will depend on the characteristics of the desired protein product, the volume needs (size of the market), and market-driven cost constraints (reviewed by Pao 1996). Transgenic plants have some remarkable features that make them particularly well suited for cost-effective bioproduction of proteins for pharmaceutical uses. These include: (a) low production costs, (b) reduced time to market, (c) unlimited supply, (d) eukaryotic protein processing, and (e) safety. Cost advantages are based not only on the low cost of biomass production, but also costs associated with research and development, greenhouse scale-up (e.g., imagine the infrastructure investment of tripling the capacity of one's aseptic fermentation or mammalian cell production facility compared to tripling one's acreage for plant growth), and reduced requirements for quality assurance testing for exclusion of human pathogenic agents (reviewed in Owen and Pao 1996). Plant-based strategies also have advantages in the pace at which feasibility testing can be done and R & D successes can be scaled up and brought to market. For example, a tobacco plant goes from seed to next generation seed in three months and produces up to a million seed per plant. Scaling-up to hundred or thousands of acres is very rapid.

Many of the therapeutic proteins of interest require complex posttranslational processing and/or oligomerization for bioactivity or appropriate targeting following administration to patients. There appears to be remarkable conservation of these protein processing steps between plants and animals such that the majority of human proteins that have been produced in plants (see Table 1) show significant structural, biochemical, and functional equivalency to proteins from humans or animal cell cultures. In cases where certain modification steps are lacking or differ in plants (e.g., glycan composition, discussed further below), strategies to introduce appropriate animal protein processing enzymes or modify the plant processing

Table 1. Feasibility of plant-based production of human (or other animal) proteins

Therapeutic product	Recombinant protein	Plant host	Structural integrity	Functional activity	Reference
Human proteins					
Human hemoglobin (H and β)	Blood substitute	Tobacco	Yes (H and β)	Yes (O ₂ -binding)	Dykes et al. 1977
Human serum albumin	Blood extender	Potato	Yes	Not tested	Singh et al. 1990
Protein C	Anticoagulant	Tobacco	Most processing steps performed	Not tested	Chen et al. 1994
Cytokines/lymphokines					
α-Interferon	Viral protection and vaccine	Rice	Yes	Yes (antiviral activity)	Zou et al. 1994
Interferon-γ	Proteoglycan	Tobacco	Yes	Yes (antiviral activity)	Chen et al. 1997
GM-CSF	Leukopenia in bone marrow transplant	Tobacco	Yes	Yes (growth stimulant of T _H 1 cells)	Dykes et al. 1996
Epidermal growth factor	Wound healing	Tobacco	CSH ^a	Not tested	Hiro et al. 1993
Tumor growth factor	Melanin	Tobacco	Yes	Not tested	Bauer et al. 1994
Erythropoietin	Malignant blood cells	Tobacco	Yes	Not tested	Martinez et al. 1993
Lysosomal enzymes	Fabry disease	Tobacco	Yes	Yes (enzyme act.)	Garcia 1997
α-Galactosidase	Glucosylceramide	Tobacco	Yes	Yes (enzyme act.)	Cassidy et al. 1998b
Other proteins					
Casein	Nutritional	Potato	CSH ^a	Not tested	Owen et al. 1997
Humulin	Anticongulant	Quinoa	Yes	Yes (clotting act.)	Panama et al. 1993
NP1 defensin	Antibiotic	Tobacco	CSH ^a	Yes (antibiotic act.)	Garcia 1997
Glucanase	Diabetes	Tobacco	CSH ^a	Yes (enzyme activity)	Ma et al. 1997

^aProtein was glycosylated but the glycan composition very differ from those produced in human

^bDetected as cross-reactive humanized animal by western blot analysis or ELISA

machinery are greatly facilitated by the ease of plant transformation and the broad experience in transgenic approaches to modifying plant metabolism through overexpression and antisense strategies. In fact, plants may be the only system capable of efficient production of certain human proteins such as growth regulators and cell cycle inhibitors which would negatively impact either the transgenic animal or animal cell culture in which they are expressed.

Perhaps the most important advantage of plants, which is emerging in the aftermath of the recent "mad cow disease" scare, involves product safety. The biopharmaceutical industry is now faced with the possibility of product validation

and quality assurance that demonstrate purity not only from known human pathogens such as HIV but also from unknown or poorly characterized agents such as the prions responsible for bovine spongiform encephalopathy and the related Creutzfeldt-Jakob diseases (Rosenau and Jakes 1996; Vaughan 1996). Plants do not serve as hosts for blood- or animal tissue-borne human pathogens. In addition, plant-based production and purification can be executed without the use of any chemical-derived products. Clearly, purity, efficacy and quality control issues similar to production of any biopharmaceutical will need to be addressed (see Myers 1997). However, plant-based bioproduction should realize significant savings as a human- and animal-source-free production system.

The list of complex human proteins and animal, viral and bacterial proteins of medical value that have been successfully expressed in plants is growing rapidly (reviewed in Owens and Peck, 1996). In addition to disease antigens (vaccines) and antibodies discussed in other chapters of this volume, transgenic plants have been used to synthesize a number of complex serum proteins, cytokines, growth regulators, anticoagulants, antitoxins, and lysosomal enzymes (see Table 1). Most of these proteins appear fully functional and structurally comparable to the native proteins produced by primary cells or in animal cell cultures or in humans. Thus, plants have clearly passed the initial test of feasibility—they are capable of producing biologic human proteins of pharmaceutical value. In addition, the first transgenic plant-synthesized products (a tobacco-derived antibody targeting gum disease and a potato-derived vaccine candidate) have reached initial human trials—a significant benchmark toward commercialization. However, as we move from feasibility studies to commercial bioproduction, issues of transgene expression levels, product processing and stability, biomass and extraction scale-up, purification, and quality control become paramount. These longer-term goals have inspired the development of novel transgene expression systems that incorporate components targeting product abundance, product recovery, and regulatory acceptance into the initial transgene design. In this review, we will discuss key issues that impact the choice and utility of plant-based production systems for biopharmaceuticals. We will highlight several production strategies that stress the importance of linking "upstream" steps in genetic engineering and expression strategies with "downstream" issues of extraction, purification, and yield. These systems are designed to separate biomass production from transgene protein production and to directly manipulate the timing, subcellular localization of the product to enhance yield, protein stability, and ease of recovery and purification.

2 Plant-Based Biopharmaceutical Production: Issues and Answers

This majority of examples demonstrating bioproduction of potential therapeutic proteins in plants shown in Table 2 have used model plant species that are easy to genetically engineer (e.g., tobacco, potato) and the "strong constitutive" 35S

Triangulo Plans for Teenage Parents: Linking Upstream and Downstream Strategies 99

Table 1. Transgene expression strategies and recombinant protein yields.

Transgene construct	Plant host	Promoter	Expression strategy, tissue	Production levels (% of soluble protein)	Reference
Secretory proteins	Tobacco	35S ^a	Secret. prot.	0.00%	Diaz et al. 1997
	Maize	35S ^a	Constitutive, leaf	0.02%	Shen et al. 1990
	Tobacco	35S	Constitutive, leaf	0.06%	Chen et al. 1996a
	Rice	P1 ^b	Constitutive, leaf	Not reported	Zhai et al. 1994
Epitopes and proteins	Tobacco	MA	Constitutive, vascularised leaf	1%	Chen 1997
	Tobacco	Rice glutelin	Secret. prot.	Not reported	Gao et al. 1996
	Tobacco	35S	Constitutive, leaf	0.001%	Hiro et al. 1993
	Tobacco	MA	Constitutive, leaf	12 mg/kg tissue	Gao 1997
Enzymes	Tobacco	MA	Peri-vascular induced leaf tissues	1%-10%	Chen et al. 1998b
	Peanut	35S ^a	Efficient	0.01%	Hao et al. 1993
	Tobacco	35S ^a	Constitutive, leaf	Not reported	Hiro et al. 1996
	Tobacco	35S ^a	Constitutive, leaf	0.097%	Mason et al. 1992
Antibodies	Onion	Oleulin	Secret. prot.	0.3% (seed protein)	Panayiotou et al. 1995
	Tobacco	35S	Constitutive	0.01%-5%	Re, in Press
	Tobacco	MA	Secret. prot.	4%-5%	Kumagai et al. 1993
	Tobacco	35S ^a	Constitutive, leaf	0.4%	Ma et al. 1997
Enzymes	Peanut	35S ^a	Efficient	0.4%	Ma et al. 1997
	Maize	35S ^a	Efficient	0.4%	Ma et al. 1997
	Maize	35S ^a	Efficient	0.4%	Ma et al. 1997
	Maize	35S ^a	Efficient	0.4%	Ma et al. 1997

Modified 3'5S promoter containing enhancer duplication and/or leader sequences (translational enhancers) from the tobacco etch virus or alfalfa mottled virus

protein derived from the calliflower mosaic virus. However, as plant biotechnology moves from demonstrating feasibility toward commercialization of protein products, many other issues come into play in selecting host species expression strategies, target tissues, and extraction/purification protocols. These choices must take into account not only the production of the particular protein of interest but

2.2 Choice of Tissue

In order to obtain maximum yields, the plant species selected must concentrate biomass in the organ or tissue where the foreign protein is expressed. The diversity among different species in this respect means that a variety of options are available including leaves, vegetative storage organs (e.g. tubers) and seeds. The tissue chosen should be compatible with the desired protein, enabling correct processing, stable accumulation and, wherever possible, efficient recovery. Many human therapeutic proteins require extensive processing for full activity, involving transport through the cellular endomembrane system. Functional lysosomal enzymes (Crawley et al. 1996b) and mammalian antibodies (Ma et al. 1993) have all been produced in leaves of tobacco following trafficking through the endoplasmic reticulum (ER) and Golgi complex. Human serum albumin has also been stably expressed in tobacco leaves and various tissues of potato including tubers (Sims et al. 1993), although the precise folding and functionality of the protein was not established. In the above examples the recombinant proteins were either specifically targeted to, and detected in, the apoplast, or presumed to locate there as a result of the default pathway of the plant endomembrane system. Deposition into the extracellular apoplast may contribute to the stability of foreign protein by removing them from the more hydrolytic intracellular environment (Pillay et al. 1993).

Expression and accumulation of foreign proteins in seeds may be achieved through compartmentalization within various subcellular storage organelles. As a natural storage organ, seeds possess attributes such as high protein content and a low hydrolytic intracellular environment that make them particularly attractive as protein production vehicles. The human neurotrophin, β -interferon (Paxinos-Narayanan et al. 1989), and the beech seed storage protein, β -amylase (Paxinos-Narayanan et al. 1993), have both been produced in seeds of *Brassica napus* following targeting to protein bodies and oil bodies respectively. Proteins can also be secreted to the apoplast of seeds. However, the recovery of apoplastic proteins from seeds may be more difficult than from some of the vegetative organs mentioned above, owing to the desiccated state of seed tissue at maturity. On the other hand, this advanced state of dehydration also confers enhanced stability, allowing seeds to be stored for periods of several years without any appreciable degradation of proteins or loss of activity (e.g. see Fox et al. 1993). The greater flexibility resulting from the separation of protein production and purification represents a distinct advantage of seeds over most other organs for which more immediate processing is often required.

2.3 Expression Strategies

Choice of promoter, which mediates the timing, tissue-specificity, and level of transgene expression, is a key determinant of transgene product yields and recovery strategies (see review by Cross and Comai 1996). As shown in Table 2, many of the human (or other animal) proteins expressed in plants have used native or enhanced

issues of recovery, purity, production/purification costs, reproducibility, supply continuity, quality control and regulatory assessment.

2.1 Selection of Crop Species

While certain features such as low production costs and high biomass capacity are common to all plant-based expression systems, other factors may strongly influence the choice of one plant species or expression strategy over another for the production of a specific foreign protein. In selecting a particular species it is important to consider how readily it can be manipulated to produce a stable transgenic line, the tissue and subcellular compartment best suited for stable expression of the heterologous protein, and the availability of methods for the efficient harvesting and initial processing of the plant material. Included in the first consideration are factors such as the amenability for transformation and regeneration of whole plants, generation time, and amenability to controlled genetic crossing. All of these factors significantly impact upon the time and resources required for product development. Plant transformation technologies are highlighted in other chapters (Hansen and Chaboo and Furr et al., this volume) and have been recently reviewed (Llewellyn 1996) and are therefore not discussed in detail here. The remaining two considerations deal mainly with product bioavailability (bioactivity, confirmation, efficacy) and recovery. Because infrastructure and methods for the harvest and processing of the major crop species already exist, wherever possible these are the species of choice. The tissue and subcellular compartment of expression determines protein processing capabilities, stability of the product and the ease with which it can be recovered.

Tobacco remains the easiest plant to genetically engineer and is widely used to test suitability of plant-based systems for bioproduction of recombinant proteins (see Table 2). Although tobacco is considered a regional crop and relatively labor intensive, at least three plant-based biotech companies are targeting tobacco for biopharmaceutical production (CropTech Corp, BioSource Technologies, Inc. and Plant Microbiology). In addition to being easily engineered, tobacco is an excellent biomass producer (in excess of 40 tons leaf fresh weight/acre based on multiple harvests per season) and prolific seed producer (up to one million seeds produced per plant), thus hastening the time in which a product can be scaled up and brought to market.

Several companies are developing production strategies involving transgene product accumulation in seeds, an organ designed to accumulate and store protein reserves (see Sect. 2.2). Companies targeting seed-based production using canola, corn or soybeans include San BioSyn Genetics, Agracetus (USA), Mogen International (the Netherlands), and Plantzyme (the Netherlands). Applied Phytologies (API, Davis, CA) is using transgenic rice and barley seed but is producing and recovering recombinant proteins during seed germination in a process analogous to mashing. Other crops being developed for biopharmaceutical protein or vaccine production include alfalfa, banana, potato, and tomato.

versions of the 35S promoter derived from the cauliflower mosaic virus to drive "constitutive" transgene expression, and it remains the most widely used promoter in plant biology for over-expression of plant proteins or inhibition via antisense strategies. The 35S promoter is active in most plant tissues (Bepko et al. 1988; Fawcett et al. 1989) and especially in its modified forms (Kay et al. 1987; Capadona and Patten 1990) can drive quite high levels of protein production. Although most of the human proteins produced using the 35S promoter (Table 2) showed accumulation levels below 0.1% of soluble protein, several transgene products (chitinase, antibodies) have been expressed at levels of 2-4-5% of extractable protein. The 35S promoter is quite active during seed development and has been used in production systems targeting recovery of recombinant proteins from seeds. However, the 35S promoter (and constitutive expression in general) has significant limitations when commercial bioproduction in massed leaves is the goal. Proteins that accumulate to high levels may negatively impact yield or overall health of the plant. High constitutive expression is sometimes associated with co-suppression or gene silencing (Taylor 1997) resulting in little or no transgene product accumulation. For proteins that are not highly stable, constitutive expression can lead to wasteful synthesis-degradation cycles and, of particular detriment for pharmaceutical application, contamination of the final product with inactive degradation products. In addition, the 35S is not highly active in many mature tissues (e.g., mature roots and fully expanded leaves) so that the full potential of biomass cannot be utilized. Use of inducible promoters or promoters that have a tight pattern of tissue- or organ-specificity avoids many of these limitations and often appears to be the strategy of choice for most companies targeting plant-based production of high-value proteins.

CropTech scientists have developed a postharvest expression system that uses an inducible promoter termed the MeGPA™ promoter (Coxeter and Wassenscheider 1997). This promoter has been modified from a defense-related gene such that it is generally inactive during normal growth and development but shows rapid and strong gene activation in response to mechanical stress (wound-induction), or mechanical gene activation) or a variety of defense stimuli. Thus, the recombinant protein is just synthesized in tobacco leaves in the field (or greenhouse). Plants can be harvested and stored for weeks in a cold room. Recombinant protein production is then induced *de novo* in the laboratory by GMP facility and fully synthesized protein recovered 8-24h later. Because survival depends on both the speed and intensity with which a plant can activate its defenses, we find the MeGPA™ promoter highly effective in driving high levels of inducible expression in all tissues of the plant including fully expanded leaves. The postharvest expression strategy has several advantages for pharmaceutical production. Biosensor production is both temporally and spatially separated from recombinant product production minimizing the impact of (a) environmental factors on protein yield and quality and (b) possible deleterious effects of transgene expression or foreign protein accumulation on plant growth and development. All recovered protein is newly synthesized. In addition, the timing of protein extraction can be adjusted based on the stability of the particular gene product to optimize yield of fully active polypep-

tides. For products requiring activation of multiple genes (e.g., multiple subunits, or target proteins that require specialized protein-modifying enzymes), construction and operation of the posturture system could also permit further encapsulation of the protein synthetases and processing machinery through addition of specific chemicals to the induction medium (e.g., inhibitors of key protein modification steps), although this could add significant expense to commercial scale bioproduction.

Bioproduction: strategies involving developmentally defined, or virally vector-mediated (e.g., Bioscience's GeneShare system) are also designed to limit recombinant protein production to a discrete period. With the GeneShare system, TMV-susceptible tobacco is field grown in an appropriate age inoculated with genetically modified virus, and harvested 2-3 weeks later for recombinant protein extraction (Gantt 1997). Within this period, the virus itself reaches high levels leading to significant transgene product accumulation. Using this system, Bioscience customers have attained very high product yield (recombinant protein representing greater than 10% of total soluble protein) and have progressed to the point of large scale seed production and pilot plant extraction. Applied Phytologics utilizes a germinator-specific promoter to direct transgene expression. Recombinant protein is produced under controlled conditions following initiation and initiation of germination of transgenic seeds, a production scheme analogous to barley malting.

Expression strategies involving seed-based accumulation of recombinant proteins may take advantage of discrete bioproduction periods and separation of transgene activity from the bulk of plant growth. A large number of seed-specific promoters, often derived from seed maturation and seed storage proteins, are available for both monocot and dicot plants. Depending on the recovery strategy (see below) and the characteristics of the protein product, promoters specific for embryonic, vegetative, and/or senescent specific expression can be selected.

2.4 Posttranslational Processing

In comparison with industrial enzyme production, bioproduction of human proteins for plant-mediated applications is particularly challenging due to the rigorous requirements with respect to purity, reproducibility, efficacy, and biocompatibility. Many of the problems with greatest promise as therapeutics require complex post-translational modifications and/or assembly. The striking fidelity with which plants appear to recognize and correctly act upon most of the processing signals employed within mammalian polypeptides indicates a high degree of conservation in protein processing machinery between plants and animals. Conserved processes include endomembranes targeting, signal peptide cleavage, protein folding and oligomerization, disulfide bond formation (although precise cysteine-cysteine bond-forming patterns have not been directly determined), asparagine-linked glycosylation, and selective retention in the ER and Golgi, and C-terminal isopeptidolysis. We have also noted internal proteolytic processing events in several human proteins expressed in tobacco that appear to mimic processing that occurs in mammalian cells

although the precise termini of the products have not yet been determined (Oishi et al. unpublished data).

However, clear differences in protein processing, most notably in glycoprotein processing, do exist between plants and animals. The glycan moiety of mammalian glycoproteins functions in protein folding and assembly, subcellular targeting, cell- or tissue-specific delivery within the body, protein half-life, and clearance from the bloodstream (Varki 1993). Thus, changes in glycan composition or arrangement are likely to affect activity or pharmacokinetics (Jensen et al. 1996; Liu 1992). Plant N-linked glycans do not contain terminal sialic acid residues as mammalian glycoproteins and contain other sugars or sugar linkages not found in mammalian glycoproteins. The same amino acid sequence (N-X-S/T) is recognized within the ER for addition of the high-mannose form glycan complex (distal) in plant and animals. However, plants process these N-linked glycans in distinct complex forms as the glycoprotein progresses through the Golgi. The sialic acid is present as the terminal sugar on many serum glycoproteins and appears to function in serum longevity and rates of clearance for tissue serum proteins (Chapman et al. 1993). Incorporation of this charged sugar residue into protein glycans has not been demonstrated in plants (Varki et al. 1993). In addition, plants do not phosphorylate high-mannose glycans—in mammals, the mannose-6-phosphate serves as a signal to target soluble glycoproteins to lysosomes. Finally, many complex plant glycans contain either fucose or xylose residues with linkages that do not occur in humans. Plant-synthesized glycoproteins displaying these sugar linkages appear highly immunogenic when injected into mammals (Chavira and Fava 1996). Interestingly, an Arabidopsis mutant defective in *N*-acetylglucosaminyl-transferase I has been identified in which all N-linked glycans are in the high-mannose form (von Sarsow et al. 1993). This report suggests that processing of glycans to complex forms is not critical for plant viability or development (in contrast to animals). Thus, plants can be altered to produce nonimmunogenic glycans. Variations in glycan composition is not unique to plant-based recombinant systems—yeast, baculovirus/insect cell, transgenic animal milk and even mammalian cell cultures often generate glycans that are heterogeneous or differ significantly from the native conformation for particular human proteins (reviewed in Jensen et al. 1996). It is clear that additional research is required for effective bioproduction of human glycoproteins in plants (discussed further for lysosomal proteins in Sect. 3). Genetic engineering strategies to modify the glycan-processing machinery of plants or *in vitro* enzymatic modification of the purified recombinant protein should enable commercialization of plant-synthesized glycoproteins for pharmaceutical applications.

Because plants are relatively easy to genetically engineer, genetic strategies to specifically alter protein processing by either antisense to block endogenous enzymes or addition of genes encoding novel processing activities are highly feasible. The recent cloning of plant genes encoding enzymes involved in Golgi-localized glycan processing opens up opportunities to modify the complex glycans produced in plants. Processes other than glycosylation can also be modified. We are interested in testing whether plants can be engineered to produce the complex serum proteinases involved in the coagulation-anticoagulation cascade (Chavez

et al. 1996a; Vaswantham et al. 1993). Plants are unlikely to perform the highly specialized γ -carboxylation of the amino-terminal glutamates required for bioreactivity of several of these enzymes (protein C, thrombin, clotting factors VII, IX and XI). We are currently introducing a human cDNA for the vitamin K dependent γ -carboxylase to perform the necessary modifications for this class of proteinase into tobacco (Chavez, Grubbs, et al. unpublished data). While these experiments are in very early stages, the concept of engineering elite plant lines for specialized protein processing for pharmaceutical bioproduction seems highly feasible.

2.5 Recovery Strategies

To capitalize on the advantages of plant-based systems in upstream production, it is necessary that downstream purification of the recombinant product be accomplished economically. Complex and inefficient purification schemes can contribute significantly to overall costs and result in lower yields so that commercial production is no longer viable. In some cases, such as in the production of individual enzymes, downstream costs can be reduced or even eliminated when a high degree of product purity is not required. A good example of this is the production of phyase in seeds. The enzyme phyase may be used to enhance the nutritional quality of seed meal by breaking down the phytates present in the meal and thereby increasing the availability of phosphate to monogastric animals. This may be conveniently achieved through expressing the phyase enzymes in seeds and adding milled transgenic seed to a standard feed meal preparation (Prest et al. 1993; Vaswantham and Fox 1996). Unfortunately, this strategy is not applicable to many proteins, particularly pharmaceutical proteins, that require rigorous purification to near-homogeneity. For these products simple and efficient methods of downstream purification must be developed.

2.5.1 Affinity Tag-Based Purification

One approach to the purification of recombinant proteins is through the use of affinity tags. This can be accomplished through the creation of a fusion between the protein of interest and another protein or peptide that exhibits affinity for a specific ligand. The fusion protein is then recovered by binding to the ligand immobilized onto a support matrix. The high selectivity possible with affinity separation often enables a substantial degree of purification to be achieved in a single step. A number of these affinity tags have been developed for use in microbial systems. Different types of ligand pairs have been exploited for this purpose including maltose binding protein-amylose, histidine residues-ni²⁺ ions, and protein A-IgG. A similar approach may be useful for the purification of recombinant proteins synthesized in plants. The efficacy of this method in plants has been demonstrated in a small scale purification of a human glucocorticoidase-FLAG epitope fusion produced in tobacco (Chavez et al. 1996b). Here, the fusion protein was recovered using an anti-FLAG antibody affinity matrix and used for bio-

from the tobacco pathogenesis-related protein, PR-5, have been used to successfully direct secretion of human serum albumin in potato (Gowans et al. 1990). Similarly, the potato proteinase inhibitor II protein signal peptide (Hizakas et al. 1995) has been used to secrete α -glucosidase into the apoplastic space of tobacco plants. While considerable enrichment of the recombinant protein can be achieved with this approach, methods of efficiently recovering proteins from the apoplastic fluid have yet to be developed.

With the appropriate signals or fusions it is also possible to target proteins to the lumen of the ER or vacuole. The human α -glucosidase *glucosylase* has been expressed in seeds of *Arabidopsis thaliana* and *Brassica napus* as an internal fusion between the N- and C-terminal ends of the Arabidopsis 2S albumin protein (Vasquez-Aguayo et al. 1995). The fusion protein was subsequently found to accumulate stably within the protein bodies of these seeds. Purification was accomplished through an initial fractionation in low salt to obtain albumin protein, followed by two proteolytic digestion steps and HPLC separation. One drawback of this strategy is the complexity of the proteolytic cleavage; particularly since carboxypeptidase was required to remove the C-terminal portion of the albumin protein. A failure to precisely control this reaction would result in significant product heterogeneity. It is also possible that folding constraints for protein body packaging might impose limitations on the size of the foreign protein that could be produced as an internal fusion.

Seed oil bodies represent another subcellular compartment available for targeting of recombinant proteins. Localization to oil bodies is achieved through creating a fusion between the desired recombinant protein and oleosin, a protein specifically targeted to these organelles. As described below, oil bodies offer some unique advantages and opportunities for expression and purification of foreign proteins.

2.5.3 Seed Oil Bodies as Purification Tools

Oil bodies are natural subcellular organelles found in all oilseeds where they form the storage site for the primary energy reserve in these seeds. Triacylglycerides (TAGs). They are composed of TAGs surrounded by a half-unit phospholipid monolayer into which is embedded a unique type of protein known as oleosin. Oleosin accumulates to high levels in oil seeds comprising between 2% and 10% of the total seed protein in different species. It is believed that the primary function of oleosin is to prevent the coalescence of oil bodies during seed desiccation. In so doing, a larger surface area is available for lipolytic enzymes enabling the rapid mobilization of TAG reserves upon seed germination. Although the protein mechanism of oleosin targeting is not fully understood, it is known that they are synthesized on the ER and that a motif in the central domain is crucial for their subsequent localization to oil bodies (Van Rooyen and Moroney 1995; Amara et al. 1997). The oleosin protein appears to consist of three distinct domains. The N- and C-terminal domains are amphipathic and proteolytic digestion studies strongly suggest that they reside on the outer surface of the oil body (Asanuma et al.

chemical studies on activity and posttranslational modifications. However, because the long-term application is as a replacement enzyme the (optimal for cellular proteins, the presence of the "nonhuman" residues is undesirable and is not used for scale-up. For some proteins and production strategies, the affinity tag can be proteolytically removed from the fusion protein following purification. However, as with any strategy involving cleavage of fusion proteins, the additional steps required to remove the tag contribute to the downstream purification costs, and there is the potential that the tag could either folding or processing of the recombinant protein.

2.5.2 Comparison of Purification

Another means of simplifying the purification of recombinant proteins is through compartmentalization. This can be achieved using either signal peptides or whole protein fusions to target the protein to a specific cellular location. In this case, purification of the desired protein is facilitated by virtue of its physical separation from other proteins in the cell. Subcellular fractionation is then used to obtain a purified fraction containing the recombinant protein. A variety of forms of compartmentalization have been recruited for the production of foreign proteins in plants. These include expression of viral particles, extracellular secretion, and targeting to intracellular storage organelles. As noted above, the posttranslational modifications required to produce a functional protein necessarily constrain where that protein can be expressed, as these reactions are, in a large extent, localized to specific subcellular compartments.

A number of plant viruses have been used for the transient expression of foreign proteins in plants (see Zetter et al. 1989; Galina, 1991; Ueda et al. 1993). To aid in purification, recombinant proteins may be engineered as fusions to viral coat proteins and then separated as mature virus particles. This strategy has recently been used in the production of malaria antigens in tobacco with tobacco mosaic virus (TMV) (Russek et al. 1995). Noninfectious plants were inoculated with infectious RNA transcribed from cDNA encoding the genetically engineered fusion protein. Mature virus particles carrying the fusion were later recovered from leaf extracts through differential centrifugation and precipitation. While in this example the purified virus particles were intended for use as a vaccine, it should also be possible to further purify recombinant proteins with this approach by introducing a protease cleavage site into the fusion protein. One possible limitation to this approach may be the size of the foreign protein, as larger proteins may impair viral coat assembly.

Secretion into the extracellular media or periplasmic space has proven to be extremely useful for production and purification of foreign proteins in many yeast and bacterial systems. In addition to providing an enriched fraction of the recombinant product, secretion has also been found to enhance protein stability and facilitate proper folding. Another attractive feature of this approach is that the signal peptide is removed from the recombinant protein in the course of normal processing enabling an authentic protein to be obtained without introducing additional proteolytic digestion steps. In plant cells, secreted proteins are deposited into the apoplastic space. The native signal peptide as well as a signal sequence

1997; Hillis et al. 1993; Tsien and Hwang 1992). The central domain is composed largely of hydrophobic amino acid residues, and is believed to adopt a hairpin conformation anchoring the protein firmly within the TAG core of the oil body. Comparison of oleosin sequences from different species reveals that the central domain is highly conserved while the N- and C-terminal exhibit considerable sequence variation.

Several features of seed oil bodies lend themselves to the production of foreign proteins. Oleosins tolerate fusion of foreign proteins to either the N- or C-terminal ends without apparent loss of oil body targeting efficiency (Morosoff and van Rooyen 1993). Oleosin fusions have been treated with a number of different proteases varying in molecular weight from approximately 7–55 kDa all of which are stably accumulated on the surface of oil bodies. In the case of the reporter enzyme β -glucuronidase, it was further shown that enzymatic activity was retained with the oleosin fusion-oil body complex. The oil bodies, together with their complement of oleosin fusion-oil body complex, are remarkably stable both within the seed and following their release by aqueous extraction (van Rooyen and Morosoff 1993b). Within the seed the proteins remain undegraded for years without the requirements for elaborate storage conditions. Following their release into aqueous solution, oil bodies are extremely resistant to mechanical disruption and are stable over a wide range in pH and temperature (Kosman et al. 1996; van Rooyen and Morosoff 1993b). Finally, the lower density of oil bodies allows them to be separated from soluble contaminants by flotation centrifugation, enabling simple and rapid purification of recombinant proteins targeted to the oil body surface. Digestion with a site-specific endoprotease to cleave the oleosin fusion protein, and centrifugation to remove the oil bodies, results in the recovery of a highly enriched fraction of the desired recombinant protein within the aqueous phase. The naturally low hydrolytic environment within the seed, coupled with the rapid removal of soluble protein contaminants, ensures that little or no degradation of the oil body-associated proteins occurs during processing. As described in Sec. 3.2, the unique properties of oleosins and oil bodies have been exploited by Sem BioSys in the development of a novel plant-based protein production and purification system.

3 Examples of Plant-synthesized Protein Therapeutics Linking Upstream and Downstream Strategies

In order to "reduce to practice" many of the considerations and strategies described above, two very different examples of plant-based bioproduction of recombinant proteins of commercial value are described below. These examples not only demonstrate the diversity of expression and purification strategies available through plants, but also highlight the convertibility of bioproduction strategies imposed by the particular protein targets. In both cases, the overall bioproduction strategy has been strongly influenced by commercial and regulatory considerations.

3.1 Production of Human Lysosomal Enzymes in *Nicotiana glauca*

Cost-effective production of recombinant human proteins for replacement enzyme therapies is likely to have a large impact on the care and treatment of patients with specific metabolic or genetic disorders. The lysosomal storage disorders represent a large class of these genetic diseases for which the molecular basis of disease has been determined and cDNAs encoding the required enzymes have been cloned (Neufeld 1991). Lysosomes, the animal organelles responsible for the regulated intracellular degradation of macromolecules, contain multiple hydrolytic including proteases, nucleases, glycosidases, lipases, phosphatases, phospholipases, and sulfatases (Davidson et al. 1988). Deficiency in specific lysosomal hydrolases can lead to toxic accumulation of the undegraded substrate and a variety of clinical manifestations. Tay-Sachs disease is perhaps the most familiar lysosomal storage disease, involving deficiencies in α -hexosaminidase that lead to accumulation of ganglioside GM₂ in the membranes of brain cells (Neufeld 1991). The mucopolysaccharidoses (MPS) are a group of lysosomal storage diseases caused by deficiencies of one or more of the ten lysosomal enzymes required for the degradation of sulfated glycosaminoglycans (reviewed in Neufeld and Morosoff 1995). Lysosomal accumulation of undegraded glycans leads to the malfunction of affected cells/organs which compromises the growth and development of the individual and may, in severe cases, lead to premature death. Replacement enzyme therapy appears promising based on human cell- and animal models, but drug development is hampered by the small patient pool and limitation in current technologies for cost-effective bioproduction. The industry paradigm for human replacement enzyme therapy is the glycoprotein product Corelase (Genzyme, Cambridge, MA) for the treatment of Gaucher disease. This lysosomal storage disorder affects 10,000–20,000 individuals in the United States (NIH Technology Assessment Panel on Gaucher Disease 1996) and is caused by defects in glucocerebrosidase, an acid β -glucosidase required for complex lipid degradation. Routine administration (generally every 2 weeks) of placental-derived enzyme has revolutionized the treatment of the disease and the quality of life of Gaucher patients. However, the high drug costs associated with purification of glucocerebrosidase from human placenta or, more recently, with bioproduction of recombinant enzyme in Chinese hamster ovary (CHO) cells, make it one of the world's most expensive drugs. Although the production of lysosomal enzymes in plants is challenging (Gardner et al. 1996b), CropTech has selected several lysosomal enzymes among its initial targets for bioproduction based on (a) the ability of plants to address critical cost, safety and supply issues for replacement enzymes, (b) the extreme medical need, and (c) the potential for Orphan Drug status to facilitate progress toward clinical trials and commercialization.

The first lysosomal enzyme produced in transgenic plants was glucocerebrosidase (EC 3.2.1.45) as a potential alternative replacement therapy for Gaucher disease (Gardner et al. 1996a,b). Placental glucocerebrosidase that has been enzymatically modified to generate mannose-terminated glycans is highly effective in

of 653 amino acids (pro-IDUA) with a signal peptidase cleavage site at amino acid 21. The cDNA for IDUA has been expressed in Cos-1 and CHO cells (Karkas et al. 1994; Scott et al. 1991) and recombinant IDUA has been purified and shown to be biologically active in dogs deficient for this enzyme (Karkas et al. 1996).

An initial test of the feasibility of commercial production of human IDUA in plants, researchers at CropTech Corporation engineered tobacco plants for postharvest production of human IDUA. The coding region from the human IDUA cDNA (Karkas et al. 1994) was placed downstream of the inducible *MeGAI* promoter and inserted into a plant transformation vector for *Agrobacterium*-mediated transformation (Horsch et al. 1983). Transgenic tobacco plants containing one to three copies of the *MeGAI* promoter-IDUA construct were tested for IDUA expression in noninduced and induced transgenic leaf material and IDUA mRNA was detected. Following 5–20% induction, abundant IDUA transcripts were detected as well as novel protein that cross-reacted with anti-IDUA antisera. Tobacco contains no endogenous IDUA activity (in contrast to the leaf β -glucosidases which are quite abundant in tobacco). Analyses of IDUA recovered from induced tobacco leaf material indicated that the IDUA protein is enzymatically active. Recombinant human IDUA requires glycosylation for activity, detection of enzyme activity in tobacco samples (as well as its electrophoretic mobility and Coomassie binding) indicates that the human IDUA signal peptide correctly targets the protein to the plant endomembrane system for glycosylation. The majority of the IDUA appears to be secreted, the default pathway for the plant endomembrane system (Demers et al. 1993). Although IDUA yields from the first IDUA-transgenic plants analyzed are lower than those seen for glucocerebrosidase-expressing plants, denaturation of enzymatic activity of the tobacco-synthesized IDUA glycoproteins and development of novel IDUA recovery methods strongly support the use of transgenic tobacco for human IDUA production.

Both glucocerebrosidase and IDUA are glycoproteins and thus pose a particular challenge for production in plants as well as other recombinant expression systems (Jaccard et al. 1990). For soluble lysosomal enzymes such as IDUA, the signal for lysosomal sorting in the mannose-6-phosphate residues present in their N-linked glycans. Mannose-6-phosphate receptors are present on the plasma membrane as well as lysosomal membranes of many mammalian cell types and thus direct uptake and lysosomal delivery of exogenously supplied IDUA (Karkas et al. 1996). Plants do not phosphorylate their glycans and glycan-based signals do not appear to function in vacuolar targeting (Fay et al. 1989; Christensen and Fay 1996). It is likely that some, if not all, of the glycans on tobacco-synthesized IDUA are in the compact form and thus likely to be immunogenic (see Sect. 2.4) and ineffective in directing the required cell-specificity for uptake and lysosomal delivery. Engineering plants to synthesize mannose-6-phosphate-modified glycans is currently not feasible – the two required enzymes have not been well characterized. However, alternative strategies that address both the delivery and immunogenicity are suggested by the currently effective lysosomal replacement therapeutic, Cerezyme. Glucocerebrosidase is a membrane-associated protein that is targeted to lysosomes by a mannose-6-phosphate-independent route. The N-linked glycans

reversing the symptoms of the disease (Barny et al. 1974). For commercial production of glucocerebrosidase (Cerezyme, Genzyme), it required between 400 and 2000 plasmids to supply a standard dose – a major factor in the extreme cost to patients (\$100,000–400,000 annually). NIH Technology Assessment Panel on Gene Therapy (1990). A CHO-synthesized recombinant form (Cerezyme, Genzyme) has been approved by the FDA, but no significant reduction in cost is anticipated. As a consequence, the success of treatment of Gaucher's disease in the United States remains limited by the cost and supply of the drug. For these reasons, it is a promising candidate for production in a plant-based system. The successful synthesis of enzymatically active human glucocerebrosidase in transgenic tobacco plants for human glucocerebrosidase was reported by Chakrabarti et al. (1996a,b). Briefly, the gene coding region for human glucocerebrosidase was modified to encode a FLAG (International Biotechnology) epitope tag for subsequent detection and purification. Aged to the inducible *MeGAI* promoter (CropTech Corporation), and introduced into tobacco. Of 49 transgenic plants analyzed, five plants had expression levels between 1%–10% total soluble leaf protein after 12h of pathogen induction. Based on biochemical kinetic studies, affinity binding characteristics, and conduction of a functional assay, the tobacco-synthesized enzyme appears fully active and comparable to human- or CHO-derived glucocerebrosidase. These findings form the basis of current and future efforts in the commercial production of glucocerebrosidase in transgenic plants for Gaucher enzyme replacement therapy. Studies to (a) identify high-expressing transgenic lines containing new glucocerebrosidase constructs that lack the nonhuman FLAG epitope and (b) test strategies to address glycan modification are now underway prior to scale-up of glucocerebrosidase production and purification technologies.

CropTech researchers have also synthesized a second lysosomal enzyme, α -mannosidase (IDUA, EC3.2.1.76), in transgenic tobacco (Jenkins, Weisenborn, Bennett, and Ogle, unpublished results). IDUA is a potential replacement therapeutic for Hunter syndrome and Hunter/Scheie syndrome, the most common MPS representing 1/100,000–1/150,000 births. Although the concept of enzyme replacement for Hunter syndrome was first investigated in the 1970s (Diffraxanra and Nicolson 1974; Nussenzweig and Mautner 1983), the development of IDUA as a drug has not progressed rapidly because of the lack of an effective production system. Recombinant enzyme sufficient for initial testing in Hunter-syndrome and feline models has been produced using a CHO-based production system (Karkas et al. 1994), but progress toward human trials is limited by protein availability. As a consequence, the successful plant-based production of IDUA has the potential to directly impact the speed of development of IDUA as a enzyme replacement therapy for Hunter and Hunter/Scheie syndromes. In humans, the lysosomal IDUA from liver is a soluble glycoprotein of 60–82kDa reflecting heterogeneity in the glycan composition. There are six potential N-linked glycosylation sites, some of which are modified to mannose-6-phosphate forms (generally sites 3 and 6) or to complex glycan forms. At all sites there is a high degree of microheterogeneity in glycans (Zhuo et al. 1996). The sequence of the complete cDNA for human IDUA has been reported (Mozzowitz et al. 1992; Scott et al. 1991) and encodes a protein

present on the placental enzyme are biorecognition structures having terminal sialic acid residues. In order to direct effective delivery to lysosomes of the affected cells in Canine patients (primarily cells of the macrophage/mesoderm lineage), sequential enzymatic digestion is used to remove the terminal sugars and expose the sialinase core (Huxton et al. 1993). This oligosaccharide form is targeted to the correct cell and organelle location to effect glycosaminoglycan degradation and symptom reduction (Quasnovas et al. 1995; Huxton et al. 1991). Complex plant glycoproteins are naturally nontoxic to humans (Quasnovas and Pavesi 1996). Enzymatic removal of the immunogenic sugar and sialic residues should yield glycoproteins of similar pharmacokinetics as Canine.

3.2 Production of Hirudin in *Brevice napus*

To evaluate the potential of Sam Brody's olefin partitioning technology, the model therapeutic protein hirudin was selected. Hirudin is a naturally occurring anticoagulant protein produced in the urinary bladder of medicinal leeches (*Hirudo medicinalis*) and is used to facilitate feeding since its discovery almost 50 years ago, it has been extensively studied. Hirudin possesses a number of desirable properties which advocate its use as a therapeutic pharmaceutical. It is an extremely specific and potent inhibitor of thrombin, the last enzyme in the blood coagulation cascade, having a K_i of 10^4 (Baker et al. 1980). It is also rapidly cleared from the body, exhibits low toxicity (500,000 IU/kg body weight in rats) (Marekwater et al. 1982) and, probably as a consequence of the covalent of leucine and methionine, has relatively low immunogenicity (Kloczko 1991). The protein has also been well characterized with respect to its structure and mechanism of binding to thrombin (Rivett et al. 1990). A small number of closely related isoforms of hirudin have been isolated all of which show strict conservation for six cysteine residues (Strong and Marekwater 1993). These residues participate in the formation of three disulfide bridges whose precise pairing is necessary for protein activity (Chambers and Quasnovas 1992, 1993). Although the native protein is isolated at the Tyr-63 position, recombinant nonaltered hirudin exhibits significant activity (Strong and Marekwater 1993). It folds spontaneously *in vitro* and functional hirudin has been produced previously in both bacterial (Harey et al. 1986; Bismar et al. 1990) and yeast (Loren et al. 1988; Loren et al. 1993) systems. However, the quantities of hirudin required were it to fully replace presently used anticoagulants such as heparin, are estimated to be on the order of hundreds to thousands of kilograms of protein annually. For this reason, hirudin is an excellent candidate for production with a high capacity plant-based system.

The common oilseed rape species, *Brassica napus*, was selected as the vehicle for production of seed-derived hirudin. After tobacco, the *Brassica* species are among those most easily transformed with *Agrobacterium*. Cells in the early cotyledonary petioles cut from young seedlings are readily infected with the bacterium. Formation of callus, regeneration to plants, and isolation of transformants are all very efficient in *B. napus*; transformation efficiencies approaching 55% of

the original explants can be obtained. The time-line for development of a transgenic plant is also relatively short, in the range of approximately 4–6 months from transformation to collection of first generation transformed seed. Another attractive feature is the availability of a highly productive system from microspore-derived embryos, facilitating the creation of homozygous lines. As an oilseed crop, considerable biomass is contained within the seed. Seed production in *B. napus* is between 1 and 2 tons per hectare at a cost of approximately (United States) \$300/ton. Protein content in these seeds represents in excess of 20% of the total seed weight, approximately 9% of which is olefin.

The production and analysis of transgenic plants expressing an olefin-hirudin fusion has been reported previously (Quasnovas et al. 1995). Briefly, a synthetic sequence encoding the hirudin variant 2 (HV2) isoform was fused to the 3' end of an *Arabidopsis* M13a olefin gene with the two coding regions separated by a sequence encoding the four amino acid recognition site for the protease, factor Xa. Following *Agrobacterium*-mediated transformation, putative transgenics were selected and expression of the olefin-hirudin fusion confirmed by northern analysis. Immunoblotting with anti-hirudin antibodies demonstrated that the olefin-hirudin fusion protein was correctly expressed and accumulated in oil bodies of transgenic seed. Oil bodies were separated and washed to remove contaminating soluble proteins through flotation-centrifugation. After digestion with factor Xa and a final round of flotation-centrifugation to remove oil bodies, hirudin was recovered in the aqueous fraction. Formation of a functional protein was confirmed by an *in vitro* thrombin inhibition assay. Comparison of protein contents in whole seed extracts and in the soluble fraction obtained after flotation-centrifugation indicated that the majority of seed protein had been removed. The enrichment obtained with this procedure demonstrates the utility of oil body compartmentalization for purification of recombinant proteins. Further purification of the recombinant hirudin to near-homogeneity was achieved through anion exchange and reverse phase chromatography. Values obtained for the specific activity of *B. napus*-derived hirudin are equivalent to those reported for recombinant hirudin produced in yeast systems (Loren et al. 1988).

3.2.1 Prospects of Olefin-Partitioning Technology

The potential for commercial application of olefin partitioning technology can be evaluated by examining the system with reference to certain key production parameters: namely, production capacity, authenticity/functionality of product, downstream purification costs, and process scalability. We have estimated the level of expression of the olefin-hirudin fusion protein in our transgenic seed to be approximately 10% of that of the endogenous olefin (Quasnovas et al. 1995). Based on this estimate, hirudin would represent approximately 0.3% of the total seed protein. While encouraging, this level is still somewhat lower than would be desired for a commercial production system. To increase expression levels we are currently testing a number of hirudin seed-specific promoters other than olefin in our fusion constructs. An increase in the expression of recombinant protein in the relatively

modest level of 1% of seed protein would result in a system capacity of approximately 2 kg of product per ton of seed. When coupled with low production costs and cost-effective purification, this level is within the range required for commercial viability.

The downstream purification of proteins synthesized as olefin fusions is greatly simplified by the oil body separation process. However, in order for this process to be cost-effective, the fusion protein cleavage step must be both efficient and economical. While useful for demonstration purposes, the factor Xa used in our initial olefin studies fails to meet these requirements. The enzyme is expensive, gives incomplete cleavage, and represented a contaminant which had to be removed in subsequent purification steps. To address this problem we are presently exploring proteases as olefin fusions immobilized on the surface of oil bodies. This will enable both economical production of the protease and easy removal following fusion protein cleavage through the existing oil body separation process. A number of suitable candidate proteases have been identified and are currently being tested.

The importance of process scale-up in determining economic feasibility is often overlooked in the initial research and development phase of a new technology. Procedures that work well for typical laboratory scale experiments cannot always be directly scaled up or easily adapted to existing industrial processes. In the case of olefin partitioning technology, we have developed and tested methods using industrial equipment for the large scale preparation of oil bodies. The results from these tests indicate that the process can be easily scaled up to meet commercial production requirements.

The recovery of active products such as lipids and β -glucanase from olefin fusions demonstrates that functional proteins can be produced using olefin partitioning technology. However, the fact that olefins are not exposed to the lumen of the ER either during synthesis or subsequent targeting to oil bodies, limits the range of different products that can be produced through olefin partitioning. Processes requiring glycosylation or other forms of posttranslational modification associated with passage through the endomembrane system would not be properly processed as olefin fusions. Nevertheless, a large number of proteins are still amenable to production using this technology. In addition to therapeutic proteins, the list includes many food and industrial enzymes. Some of these products are presently under development. Additionally, the ability to produce functional proteins on the surface of oil bodies offers exciting new possibilities for the production of immobilized protein matrices (Gross et al. 1996). With continued development in each of the areas mentioned above, prospects for the successful commercialization of olefin partitioning technology appear very promising.

4 Summary

We have described two very different and innovative plant-based production systems - postharvest production and recovery of recombinant product from tobacco

leaves using an inducible promoter and olefin-mediated recovery of recombinant product from oilseeds using a seed-specific promoter. Both basic technologies are broadly applicable to numerous classes of pharmaceutical and industrial proteins. As with any emerging technology, the key to success may lie in identifying those products and applications that would most benefit from the unique advantages offered by each system. The postharvest tobacco leaf system appears effective for proteins requiring complex posttranslational processing and endomembrane targeting. Because of the remarkable fecundity and biomass production capacity of tobacco, biomass scale-up is very rapid and production costs are low. Clearly this development of a quickly cost-effective extraction and purification technology will be critical for full realization of the commercial opportunities afforded by transgenic plant-based bioproduction. The recovery of protein from tobacco leaves or olefin partitioned proteins by oil-body separation represents significant breakthrough for cost-effective commercialization strategies. Additional low-cost, high-affinity separation technologies need to be developed for effective scale-up production of plant-synthesized recombinant proteins. Clearly successful commercialization of plant-synthesized biopharmaceuticals must effectively link upstream strategies involving gene and protein design with downstream strategies for reproducible GMP-level recovery of bioactive recombinant protein. Both the tobacco and oilseed systems are uniquely designed to address issues of biomass storage, product recovery, quality assurance and regulatory scrutiny in addition to issues of transgenic expression and protein processing.

References

- Abel RM, Hoberart JA, Albers M, Murphy M, Mocher JDA (1997) Role of the protein fold itself in olefin-mediated protein targeting and oil body targeting. *Plant Cell* 9:141-149
- Bevan MW, Brady RO, Drenthwaite JM, Simonsen AM, Russell RS, Hill SC, Martin BJ, Murray GA, Parker PJ, Arnold CE, Crowell HP, Yu KY (1991) Endometrial therapy of advanced cancer: oil body-mediated targeted phenolamide delivery for ovarian cancer. *N Engl J Med* 324:1464
- Bevan M, Vogel R, Kralic R, Palmer J (1996) Synthesis and secretion of lipids by *Arabidopsis thaliana*. *Plant Physiol* 110:303-307
- Bevan MW, Long R, Chen W (1997) The CAVY 303 oilseed contains at least two distinct oil bodies with different developmental and tissue-specific expression patterns. *Plant Cell* 9:2707-2722
- Bevan M, Seed B, Knepper G (1996) A novel protein hormone is expressed, correctly folded and partially processed in the leaves but not the seed of transgenic tobacco. *Transgenic Res* 5:305-310
- Berry RD, Peniston PG, Gal AE, Hilbert BR, Dandekar AS (1976) Replacement therapy for inherited enzyme deficiency: Use of purified glucocerebrosidase in Gaucher's disease. *N Engl J Med* 29:1598-1603
- Brown PJ, Danti S, McWhirter J, Shaw OR (1983) Use of site-directed mutagenesis to investigate the basis for the specificity of lipase. *Biochemistry* 22:5371-5377
- Chatterjee JC, Freed DD (1996) Carboxyl-terminal replacement of translation by a plant polyprotein. *Proc Natl Acad Sci USA* 93:1590-1597
- Chatterjee JC, Chong J (1993) The folding of lipase: a mechanism of fold and error. *J Biol Chem* 268:3008-3014
- Chatterjee JC, Chong J (1993) The olefin-folding pathway of lipase: elucidated by olefin-folding experiments. *J Biol Chem* 268:2985-2996

- Hogg E, Kato Y, Aoki H (1993) Expression of a cationically synthesized gene for human cytochrome P-450 3A4 in transgenic mice. *Biochem Biophys Res Commun* 197:1477-1481.
- Hilli MA, Wallace MD, Murphy DJ (1993) Targeting of elements to the oil bodies of oilseed rape *Brassica napus*. *Plant Mol Biol* 21:67-74.
- Hornes KR, Fry JR, Hoffmann NT, Eisenblatt D, Rogers SD, Fealey RT (1985) A simple and general method for transferring genes into plants. *Science* 227:1225-1231.
- Jacobson NJ, Knack RZ, Jones DC (1986) Getting the glycoprotein right: implications for the biotechnology industry. *Nat Biotechnol* 4:975-981.
- Kay R, Cima A, Tsai M, McPherson J (1987) Duplication of ODAV 335 promoter sequence creates a strong enhancer for plant genes. *Science* 234:1299-1302.
- Kubitzki ED, Mayfield AC, Jones AL, Neill RD (1989) Overexpression of the human lysozymal enzyme as a transgene in Chinese hamster ovary cells. *Proc Soc Exp Biol Med* 192:225-232.
- Kuttsen ED, Melchers MP, Rosenblum A, Verhoff AD, Ward DA, Oetting RA, Kanis S, Boddin C, Chien SL, Schulz RM (1993) Long-term oral high-dose trial of enzyme replacement therapy in the canine model of mucopolysaccharidosis I. *Biochem Mol Med* 58:158-167.
- Kubitzki ED (1993) Technology of transgenic animals. In: Tizard IR (ed) *Transgenic Animals*, pp 122-129. Marcel Dekker, New York.
- Lambert M, Vargiu LA, Molteni MA, van den Broek PH (1996) Oil bodies of transgenic *Arabidopsis thaliana* express functional β -glucuronidase. *J Am Oil Chem Soc* 73:1533-1538.
- Lammi MJ, Thompson TJ, Weizman N, Della-Croce D, Torres AM, Danoso J, Hill ME, Grunbaum AG, Dawson WO, Chow ET, Padak ML, GELIX (1993) Rapid high-level expression of heterologous active recombinant proteins by an RNA virus vector. *Proc Natl Acad Sci USA* 90:477-480.
- Latham RW, Joyce JO, Bailey FR, Michalski EJ, Smith LS, Davis JC, Jacobson MA, Miller WJ (1993) Expression, purification and characterization of multigene systems of a recombinant by wild-type β -D-glucuronidase variant expressed in Saccharomyces cerevisiae. *Protein Expr Purif* 6:247-255.
- Lindsay JC (1992) Plant transformation systems. In: Owen MRL, Pen J (eds) *Transgenic Plants: A Practical Guide*, pp 5-26. Academic Press, London.
- Levine G, Gludovsk A, Bernard S, Nguyen-Jaimes M, Marquet M, Bish-Shabun N, Cavalli D, O'Connor S, Siddle K, Brown SW, Courtney M, Reichel C, Lemley V (1988) Expression and secretion in *S. cerevisiae* of heterologously active leech hirudin. *Biochemistry* 27:72-77.
- Lui STY (1992) Organoplastic pharmaceuticals—science and regulatory considerations and the United States Drug Act. *Trends Biotechnol* 10:114-120.
- Maj RK, Choi A, Vire NO, Ward P, Smith P, van DerWoude C, Webster K, Lehrer T (1993) Production and assembly of secretory antibodies in plants. *Science* 260:716-718.
- Mao S, Zhao Z, Yin A, Mickelson R, Shih B, Qiu H, Solter CL, Benham ASM (1992) Transgenic plants expressing recombinant IgG to raise human and murine responses. *Nat Med* 1:753-756.
- Markwardt F, Hauptmann U, Nurnst G, Klein G, Waltemeyer P (1992) Pharmacological studies on the antitumor activity of fumagillin in experimental models. *Toxicol Hematol* 47:226-233.
- Mason HS, Lam D, Anderson CJ (1992) Expression of hepatitis B surface antigen in transgenic plants. *Proc Natl Acad Sci USA* 89:7194-7197.
- McIntosh S, Threlkeld K, Ueda M, Smith L (1993) Characterization of a human glycoprotein (erythropoietin) produced in cultured tobacco cells. *Plant Mol Biol* 21:163-172.
- Milde J (1997) Plants as bioreactors for biopharmaceutical regulatory applications. *Trends Biotechnol* 15:45-50.
- Mohoney MM and van Soest GH (1996) Recombinant protein via oleate partitioning before 7:105-113.
- Mohtashemi AS, Diani B, Chuang PD, Nardelli BP (1992) Cloning and expression of a cDNA encoding the human lysosomal enzyme α -L-fucosidase. *FASEB J* 6:677 (abstract).
- NIR Technology Associates Inc Patent on Genetically Engineered Tobacco (1990) Genetic disease: current issues in diagnosis and treatment. *JAMA* 273:948-953.
- Neill RD (1991) Lysosomal storage diseases. *Annu Rev Biochem* 60:237-250.
- Nordberg KF, Atanasiu J (1989) The mucopolysaccharidoses. In: Berrier CR, Beaudet AL, Byrnes WE, Valle D (eds) *The metabolic basis of inherited disease* 6th edn. McGraw-Hill, New York, pp 1553-1587.
- Nordberg KF, Nordberg J (1993) The mucopolysaccharidoses. In: Scriver CR, Beaudet AL, Sly WS, Valle D (eds) *The metabolic and molecular bases of inherited disease* 7th edn. McGraw-Hill, New York, pp 2463-2494.
- Owen MRL, Pen J (1996) Transgenic plants: a production system for industrial and pharmaceutical molecules. Wiley, Chichester.
- Ortiz CM, Camarero F, Ortiz-Sanchez J, Ortiz-Molina E, Pen J (eds) *Transgenic plants: a production system for industrial and pharmaceutical products*. Wiley, Chichester, pp 31-47.
- Peng QM, DAX, Roberts W, Johnson T, Wu X, Bai G, Slattery CH, Lindgren WHR (1997) Expression of the human cystic fibrosis Δ -F508 transgene in transgenic potato plants. *Transgenic Res* 6:269-286.
- Chapman MS, Frye J (1996) The production of recombinant glycoproteins with defined two-dimensional glycosylation in *Oenothera*. *Plant Cell Rep* 15:109-114.
- Cramer CL, Weinberger DL, Old KB, Graham BA, Bennett S, Freed R, Graesslin GA, Radin DN (1996) Bioproduction of human enzymes in transgenic tobacco. In: Collins GE, Shepherd RJ (eds) *Engineering Plants for Commercial Products and Applications*. New York: Academy of Sciences, New York, pp 6-7.
- Cramer CL, Weinshenker DL, Old KB, Radin DN (1996) High-level production of extracellularly active human lysozymal protein in transgenic plants. In: Owen MRL, Pen J (eds) *Transgenic plants: a production system for industrial and pharmaceutical products*. Wiley, Chichester, pp 299-310.
- Cramer CL, Weinshenker DL (1997) HMG2 promotes expression levels and post-transcriptional processing in plants. *Plant Cell Rep* 16:307-310.
- Quercia F, Leclerc H, Dubois P (1993) Molecular cell biology. Academic Press, San Diego.
- Dawson J, Steinman J, Robinson D (1993) Protein secretion in plant cells can occur via a default pathway. *Plant Cell* 5:21-31.
- De Luca CA, Finkbeiner JR, Horvath GS, Ma AI, P. Schultze M, Spohn T (1989) The expression, localization and effect of a human interferon in plants. *Virology* 172:313-322.
- Dijkwel P, Peleman J, Feyens C, Marcks MC, Gaster W, Toimati P, Buitendijck A, Menot B (1997) Human immunoglobulin-like transgenic tobacco. *Nat Biotech* 15:32-36.
- Dijkwel P, Marcks MC, Niekamp RL Jr (1994) Comments on the plasma treatment of the mucopolysaccharide. *Birth Defects* 30:214-215.
- Duke MK, Nay P, Srivastava SK, Chua NH (1993) Multiple over-regulatory demand for maximal expression of the antifreeze protein from transgenic tobacco. *Plant Cell* 5:143-150.
- Frye JE, Johnson ED, Simon A, Chelappadurai M (1993) Tobacco, biotechnology, and function of agricultural crops. *Plant Physiol* 123:303-316.
- Gilbert J, Dwyer J, Owen MRL, Gundersen A, Cockburn A, Williams OC (1993) Secretion of a functional single-chain Fc protein in transgenic tobacco plants and cell suspension cultures. *Plant Mol Biol* 23:61-69.
- Grove PL, Davies AK, Tebbey EH, Seiden R, Sander C, Cheng X, Altman T (1994) Depression of human blood proteins in transgenic plants: the cytokine GM-CSF as a model protein. In: Owen MRL, Pen J (eds) *Transgenic plants: a production system for industrial and pharmaceutical products*. Wiley, Chichester, pp 281-297.
- Harlow RB, Galloway GA, Barton NW, Fujiwara G, Damstra JJ, Banquet TK, McKee MA, Parker C, Schiffman S, Hill SC, Brady RO (1992) Reagent therapy in type I Gaucher disease: comparative efficacy of marrow-transplanted glomerular cells from normal and recombinant sources. *Ann Intern Med* 122:31-39.
- Hendell LR (1993) Tobacco mosaic virus as a gene expression vector. *Agro-Food Industry Biotech News* December: 27-29.
- Hill KL (1997) First revised, large scale production of drugs and pharmaceuticals in plants. Presentation at BEC 1st Annual International Symposium on Producing the Next Generation of Therapeutic Drugs Using Transgenic Technologies, West Palm Beach, 3-5 Feb.
- Johnson ED, Wells JD, Gerber B (1991) Glycosylation of human proteins C₁ esterase inhibitor, preventing Ascaris infections and activated by thrombin. *J Biol Chem* 266:9779-9783.
- Kuo TC, Wang FS, Chenotsky DI, Anderson CJ (1993) Oral immunization with a recombinant bacterial surface protein in transgenic plants. *Science* 262:714-716.
- Kuo TC, Dwyer B, Stokell L, Schreiber F, Courtney P, Courtney M, Tolchower R, Loong JP (1996) Cloning and expression of cDNA coding for anticapsule protein from blood-sucking leech, *Manduca sexta*. *Proc Natl Acad Sci USA* 93:1041-1046.
- Collins GE, Shepherd RJ (eds) *Engineering plants for commercial products and applications*. New York: Academy of Sciences, New York, pp 36-56.
- Gruber K, Wilke S, Sommerfeld U (1993) A thermolabile xylanase from *Clavibacter dermatophilus* expressed at high levels in the tissues of transgenic tobacco has no detrimental effects and is easily purified. *Biochemistry* 32:65-68.

11. C.L. Granger et al., *Transcatheter Platelets for Thrombotic Protection*

- [illegible]